Exhibit 2

"The public conversation about surveillance in the digital age would be a good deal more intelligent if we all read Bruce Schneier first."

—MALCOLM GLADWELL

DATA AND GOLIATH

The Hidden Battles to Collect Your Data and Control Your World

BRUCE SCHNEIER

DATA AND GOLIATH

The Hidden Battles to Collect Your Data and Control Your World

BRUCE SCHNEIER

W. W. NORTON & COMPANY
NEW YORK • LONDON

producing intimate personal data about you. It includes what you read, watch, and listen to. It includes whom you talk to and what you say. Ultimately, it covers what you're thinking about, at least to the extent that your thoughts lead you to the Internet and search engines. We are living in the golden age of surveillance.

Sun Microsystems' CEO Scott McNealy said it plainly way back in 1999: "You have zero privacy anyway. Get over it." He's wrong about how we should react to surveillance, of course, but he's right that it's becoming harder and harder to avoid surveillance and maintain privacy.

Surveillance is a politically and emotionally loaded term, but I use it deliberately. The US military defines surveillance as "systematic observation." As I'll explain, modern-day electronic surveillance is exactly that. We're all open books to both governments and corporations; their ability to peer into our collective personal lives is greater than it has ever been before.

The bargain you make, again and again, with various companies is surveillance in exchange for free service. Google's chairman Eric Schmidt and its director of ideas Jared Cohen laid it out in their 2013 book, *The New Digital Age*. Here I'm paraphrasing their message: if you let us have all your data, we will show you advertisements you want to see and we'll throw in free web search, e-mail, and all sorts of other services. It's convenience, basically. We are social animals, and there's nothing more powerful or rewarding than communicating with other people. Digital means have become the easiest and quickest way to communicate. And why do we allow governments access? Because we fear the terrorists, fear the strangers abducting our children, fear the drug dealers, fear whatever bad guy is in vogue at the moment. That's the NSA's justification for its mass-surveillance programs; if you let us have all of your data, we'll relieve your fear.

The problem is that these aren't good or fair bargains, at least as they're structured today. We've been accepting them too easily, and without really understanding the terms.

Here is what's true. Today's technology gives governments and corporations robust capabilities for mass surveillance. Mass surveillance is dangerous. It enables discrimination based on almost any criteria: race, religion, class, political beliefs. It is being used to control what we see, what we can do, and, ultimately, what we say. It is being done without offering citizens recourse or any real ability to opt out, and without any meaningful checks and balances. It makes us less safe. It makes us less free. The rules we had established to protect us from these dangers under earlier technological regimes are now woefully insufficient; they are not working. We need to fix that, and we need to do it very soon.

In this book, I make that case in three parts.

Part One describes the surveillance society we're living in. Chapter 1 looks at the varieties of personal data we generate as we go about our lives. It's not just the cell phone location data I've described. It's also data about our phone calls, e-mails, and text messages, plus all the webpages we read, our financial transaction data, and much more. Most of us don't realize the degree to which computers are integrated into everything we

Chapter 8 turns to the harms caused by unfettered corporate surveillance. Private companies now control the "places" on the Internet where we gather, and they're mining the information we leave there for their own benefit. By allowing companies to know everything about us, we're permitting them to categorize and manipulate us. This manipulation is largely hidden and unregulated, and will become more effective as technology improves.

Ubiquitous surveillance leads to other harms as well. Chapter 9 discusses the economic harms, primarily to US businesses, that arise when the citizens of different countries try to defend themselves against surveillance by the NSA and its allies. The Internet is a global platform, and attempts by countries like Germany and Brazil to build national walls around their data will cost companies that permit government surveillance —particularly American companies—considerably.

In Chapter 10, I discuss the harms caused by a loss of privacy. Defenders of surveillance—from the Stasi of the German Democratic Republic to the Chilean dictator Augusto Pinochet to Google's Eric Schmidt—have always relied on the old saw "If you have nothing to hide, then you have nothing to fear." This is a dangerously narrow conception of the value of privacy. Privacy is an essential human need, and central to our ability to control how we relate to the world. Being stripped of privacy is fundamentally dehumanizing, and it makes no difference whether the surveillance is conducted by an undercover policeman following us around or by a computer algorithm tracking our every move.

In Chapter 11, I turn to the harms to security caused by surveillance. Government mass surveillance is often portrayed as a security benefit, something that protects us from terrorism. Yet there's no actual proof of any real successes against terrorism as a result of mass surveillance, and significant evidence of harm. Enabling ubiquitous mass surveillance requires maintaining an insecure Internet, which makes us all less safe from rival governments, criminals, and hackers.

Finally, Part Three outlines what we need to do to protect ourselves from government and corporate surveillance. The remedies are as complicated as the issues, and often require fine attention to detail. Before I delve into specific technical and policy recommendations, though, Chapter 12 offers eight general principles that should guide our thinking.

The following two chapters lay out specific policy recommendations: for governments in Chapter 13, and for corporations in Chapter 14. Some of these recommendations are more detailed than others, and some are aspirational rather than immediately implementable. All are important, though, and any omissions could subvert the other solutions.

Chapter 15 turns to what each of us can do individually. I offer some practical technical advice, as well as suggestions for political action. We're living in a world where technology can trump politics, and also where politics can trump technology. We need both to work together.

1

Data as a By-product of Computing

Computers constantly produce data. It's their input and output, but it's also a by-product of everything they do. In the normal course of their operations, computers continuously document what they're doing. They sense and record more than you're aware of.

For instance, your word processor keeps a record of what you've written, including your drafts and changes. When you hit "save," your word processor records the new version, but your computer doesn't erase the old versions until it needs the disk space for something else. Your word processor automatically saves your document every so often; Microsoft Word saves mine every 20 minutes. Word also keeps a record of who created the document, and often of who else worked on it.

Connect to the Internet, and the data you produce multiplies: records of websites you visit, ads you click on, words you type. Your computer, the sites you visit, and the computers in the network each produce data. Your browser sends data to websites about what software you have, when it was installed, what features you've enabled, and so on. In many cases, this data is enough to uniquely identify your computer.

Increasingly we communicate with our family, friends, co-workers, and casual acquaintances via computers, using e-mail, text messaging, Facebook, Twitter, Instagram, SnapChat, WhatsApp, and whatever else is hot right now. Data is a by-product of this high-tech socialization. These systems don't just transfer data; they also create data records of your interactions with others.

Walking around outside, you might not think that you're producing data, but you are. Your cell phone is constantly calculating its location based on which cell towers it's near. It's not that your cell phone company particularly cares where you are, but it needs to know where your cell phone is to route telephone calls to you.

Of course, if you actually use that phone, you produce more data: numbers dialed and calls received, text messages sent and received, call duration, and so on. If it's a smartphone, it's also a computer, and all your apps produce data when you use them—and sometimes even when you're not using them. Your phone probably has a GPS receiver, which produces even more accurate location information than the cell tower location alone. The GPS receiver in your smartphone pinpoints you to within 16 to 27 feet; cell towers, to about 2,000 feet.

Purchase something in a store, and you produce more data. The cash register is a computer, and it creates a record of what you purchased and the time and date you purchased it. That data flows into the merchant's computer system. Unless you paid cash, your credit card or debit card information is tied to that purchase. That data is also sent to the credit card company, and some of it comes back to you in your monthly bill.

There may be a video camera in the store, installed to record evidence in case of theft or fraud. There's another camera recording you when you use an ATM. There are more cameras outside, monitoring buildings, sidewalks, roadways, and other public spaces.

Get into a car, and you generate yet more data. Modern cars are loaded with computers, producing data on your speed, how hard you're pressing on the pedals, what position the steering wheel is in, and more. Much of that is automatically recorded in a black box recorder, useful for figuring out what happened in an accident. There's even a computer in each tire, gathering pressure data. Take your car into the shop, and the first thing the mechanic will do is access all that data to diagnose any problems. A self-driving car could produce a gigabyte of data per second.

Snap a photo, and you're at it again. Embedded in digital photos is information such as the date, time, and location—yes, many cameras have GPS—of the photo's capture; generic information about the camera, lens, and settings; and an ID number of the camera itself. If you upload the photo to the web, that information often remains attached to the file.

It wasn't always like this. In the era of newspapers, radio, and television, we received information, but no record of the event was created. Now we get our news and entertainment over the Internet. We used to speak to people face-to-face and then by telephone; we now have conversations over text or e-mail. We used to buy things with cash at a store; now we use credit cards over the Internet. We used to pay with coins at a tollbooth, subway turnstile, or parking meter. Now we use automatic payment systems, such as EZPass, that are connected to our license plate number and credit card. Taxis used to be cash-only. Then we started paying by credit card. Now we're using our smartphones to access networked taxi systems like Uber and Lyft, which produce data records of the transaction, plus our pickup and drop-off locations. With a few specific exceptions, computers are now everywhere we engage in commerce and most places we engage with our friends.

Last year, when my refrigerator broke, the serviceman replaced the computer that controls it. I realized that I had been thinking about the refrigerator backwards: it's not a refrigerator with a computer, it's a computer that keeps food cold. Just like that, everything is turning into a computer. Your phone is a computer that makes calls. Your car is a computer with wheels and an engine. Your oven is a computer that bakes lasagnas. Your camera is a computer that takes pictures. Even our pets and livestock are now regularly chipped; my cat is practically a computer that sleeps in the sun all day.

Computers are getting embedded into more and more kinds of products that connect to the Internet. A company called Nest, which Google purchased in 2014 for more than $3 billion, makes an Internet-enabled thermostat. The smart thermostat adapts to your behavior patterns and responds to what's happening on the power grid. But to do all that, it records more than your energy usage: it also tracks and records your home's temperature, humidity, ambient light, and any nearby movement. You can buy a smart refrigerator that tracks the expiration dates of food, and a smart air conditioner that can learn your preferences and maximize energy efficiency. There's more coming: Nest is now selling a

smart smoke and carbon monoxide detector and is planning a whole line of additional home sensors. Lots of other companies are working on a wide range of smart appliances. This will all be necessary if we want to build the smart power grid, which will reduce energy use and greenhouse gas emissions.

We're starting to collect and analyze data about our bodies as a means of improving our health and well-being. If you wear a fitness tracking device like Fitbit or Jawbone, it collects information about your movements awake and asleep, and uses that to analyze both your exercise and sleep habits. It can determine when you're having sex. Give the device more information about yourself—how much you weigh, what you eat—and you can learn even more. All of this data you share is available online, of course.

Many medical devices are starting to be Internet-enabled, collecting and reporting a variety of biometric data. There are already—or will be soon—devices that continually measure our vital signs, our moods, and our brain activity. It's not just specialized devices; current smartphones have some pretty sensitive motion sensors. As the price of DNA sequencing continues to drop, more of us are signing up to generate and analyze our own genetic data. Companies like 23andMe hope to use genomic data from their customers to find genes associated with disease, leading to new and highly profitable cures. They're also talking about personalized marketing, and insurance companies may someday buy their data to make business decisions.

Perhaps the extreme in the data-generating-self trend is lifelogging: continuously capturing personal data. Already you can install lifelogging apps that record your activities on your smartphone, such as when you talk to friends, play games, watch movies, and so on. But this is just a shadow of what lifelogging will become. In the future, it will include a video record. Google Glass is the first wearable device that has this potential, but others are not far behind.

These are examples of the Internet of Things. Environmental sensors will detect pollution levels. Smart inventory and control systems will reduce waste and save money. Internet-connected computers will be in everything—smart cities, smart toothbrushes, smart lightbulbs, smart sidewalk squares, smart pill bottles, smart clothing—because why not? Estimates put the current number of Internet-connected devices at 10 billion. That's already more than the number of people on the planet, and I've seen predictions that it will reach 30 billion by 2020. The hype level is pretty high, and we don't yet know which applications will work and which will be duds. What we do know is that they're all going to produce data, lots of data. The things around us will become the eyes and ears of the Internet.

The privacy implications of all this connectivity are profound. All those smart appliances will reduce greenhouse gas emissions—and they'll also stream data about how people move around within their houses and how they spend their time. Smart streetlights will gather data on people's movements outside. Cameras will only get better, smaller, and more mobile. Raytheon is planning to fly a blimp over Washington, DC, and Baltimore in 2015 to test its ability to track "targets"—presumably vehicles—on the ground, in the water, and in the air.

The upshot is that we interact with hundreds of computers every day, and soon it will be thousands. Every one of those computers produces data. Very little of it is the obviously juicy kind: what we ordered at a restaurant, our heart rate during our evening jog, or the last love letter we wrote. Rather, much of it is a type of data called *metadata*. This is data about data—information a computer system uses to operate or data that's a by-product of that operation. In a text message system, the messages themselves are data, but the accounts that sent and received the message, and the date and time of the message, are all metadata. An e-mail system is similar: the text of the e-mail is data, but the sender, receiver, routing data, and message size are all metadata—and we can argue about how to classify the subject line. In a photograph, the image is data; the date and time, camera settings, camera serial number, and GPS coordinates of the photo are metadata. Metadata may sound uninteresting, but, as I'll explain, it's anything but.

Still, this smog of data we produce is not necessarily a result of deviousness on anyone's part. Most of it is simply a natural by-product of computing. This is just the way technology works right now. Data is the exhaust of the information age.

HOW MUCH DATA?

Some quick math. Your laptop probably has a 500-gigabyte hard drive. That big backup drive you might have purchased with it can probably store two or three terabytes. Your corporate network might have one thousand times that: a petabyte. There are names for bigger numbers. A thousand petabytes is an exabyte (a billion billion bytes), a thousand exabytes is a zettabyte, and a thousand zettabytes is a yottabyte. To put it in human terms, an exabyte of data is 500 billion pages of text.

All of our data exhaust adds up. By 2010, we as a species were creating more data per day than we did from the beginning of time until 2003. By 2015, 76 exabytes of data will travel across the Internet every year.

As we start thinking of all this data, it's easy to dismiss concerns about its retention and use based on the assumption that there's simply too much of it to save, and in any case it would be too hard to sift through for nuggets of meaningful information. This used to be true. In the early days of computing, most of this data—and certainly most of the metadata—was thrown away soon after it was created. Saving it took too much memory. But the cost of all aspects of computing has continuously fallen over the years, and amounts of data that were impractical to store and process a decade ago are easy to deal with today. In 2015, a petabyte of cloud storage will cost $100,000 per year, down 90% from $1 million in 2011. The result is that more and more data is being stored.

You could probably store every tweet ever sent on your home computer's disk drive. Storing the voice conversation from every phone call made in the US requires less than 300 petabytes, or $30 million, per year. A continuous video lifelogger would require 700 gigabytes per year per person. Multiply that by the US population and you get 2 exabytes per year, at a current cost of $200 million. That's expensive but plausible, and the price will only go down. In 2013, the NSA completed its massive Utah Data Center in Bluffdale. It's currently the third largest in the world, and the first of several that the NSA

2

Data as Surveillance

Governments and corporations gather, store, and analyze the tremendous amount of data we chuff out as we move through our digitized lives. Often this is without our knowledge, and typically without our consent. Based on this data, they draw conclusions about us that we might disagree with or object to, and that can impact our lives in profound ways. We may not like to admit it, but we are under mass surveillance.

Much of what we know about the NSA's surveillance comes from Edward Snowden, although people both before and after him also leaked agency secrets. As an NSA contractor, Snowden collected tens of thousands of documents describing many of the NSA's surveillance activities. In 2013, he fled to Hong Kong and gave them to select reporters. For a while I worked with Glenn Greenwald and the *Guardian* newspaper, helping analyze some of the more technical documents.

The first news story to break that was based on the Snowden documents described how the NSA collects the cell phone call records of every American. One government defense, and a sound bite repeated ever since, is that the data collected is "only metadata." The intended point was that the NSA wasn't collecting the words we spoke during our phone conversations, only the phone numbers of the two parties, and the date, time, and duration of the call. This seemed to mollify many people, but it shouldn't have. Collecting metadata on people means putting them under surveillance.

An easy thought experiment demonstrates this. Imagine that you hired a private detective to eavesdrop on someone. The detective would plant bugs in that person's home, office, and car. He would eavesdrop on that person's phone and computer. And you would get a report detailing that person's conversations.

Now imagine that you asked the detective to put that person under surveillance. You would get a different but nevertheless comprehensive report: where he went, what he did, who he spoke with and for how long, who he wrote to, what he read, and what he purchased. That's metadata.

Eavesdropping gets you the conversations; surveillance gets you everything else.

Telephone metadata alone reveals a lot about us. The timing, length, and frequency of our conversations reveal our relationships with others: our intimate friends, business associates, and everyone in-between. Phone metadata reveals what and who we're interested in and what's important to us, no matter how private. It provides a window into our personalities. It yields a detailed summary of what's happening to us at any point in time.

A Stanford University experiment examined the phone metadata of about 500 volunteers over several months. The personal nature of what the researchers could deduce

from the metadata surprised even them, and the report is worth quoting:

- Participant A communicated with multiple local neurology groups, a specialty pharmacy, a rare-condition management service, and a hotline for a pharmaceutical used solely to treat relapsing multiple sclerosis.
- Participant B spoke at length with cardiologists at a major medical center, talked briefly with a medical laboratory, received calls from a pharmacy, and placed short calls to a home reporting hotline for a medical device used to monitor cardiac arrhythmias.
- Participant C made a number of calls to a firearms store that specializes in the AR semiautomatic rifle platform, and also spoke at length with customer service for a firearm manufacturer that produces an AR line.
- In a span of three weeks, Participant D contacted a home improvement store, locksmiths, a hydroponics dealer, and a head shop.
- Participant E had a long early morning call with her sister. Two days later, she placed a series of calls to the local Planned Parenthood location. She placed brief additional calls two weeks later, and made a final call a month after.

That's a multiple sclerosis sufferer, a heart attack victim, a semiautomatic weapons owner, a home marijuana grower, and someone who had an abortion, all from a single stream of metadata.

Web search data is another source of intimate information that can be used for surveillance. (You can argue whether this is data or metadata. The NSA claims it's metadata because your search terms are embedded in the URLs.) We don't lie to our search engine. We're more intimate with it than with our friends, lovers, or family members. We always tell it exactly what we're thinking about, in words as clear as possible. Google knows what kind of porn each of us searches for, which old lovers we still think about, our shames, our concerns, and our secrets. If Google decided to, it could figure out which of us is worried about our mental health, thinking about tax evasion, or planning to protest a particular government policy. I used to say that Google knows more about what I'm thinking of than my wife does. But that doesn't go far enough. Google knows more about what I'm thinking of *than I do*, because Google remembers all of it perfectly and forever.

I did a quick experiment with Google's autocomplete feature. This is the feature that offers to finish typing your search queries in real time, based on what other people have typed. When I typed "should I tell my w," Google suggested "should i tell my wife i had an affair" and "should i tell my work about dui" as the most popular completions. Google knows who clicked on those completions, and everything else they ever searched on.

Google's CEO Eric Schmidt admitted as much in 2010: "We know where you are. We know where you've been. We can more or less know what you're thinking about."

If you have a Gmail account, you can check for yourself. You can look at your search history for any time you were logged in. It goes back for as long as you've had the account, probably for years. Do it; you'll be surprised. It's more intimate than if you'd sent Google your diary. And even though Google lets you modify your ad preferences, you have no rights to delete anything you don't want there.

There are other sources of intimate data and metadata. Records of your purchasing habits reveal a lot about who you are. Your tweets tell the world what time you wake up in the morning, and what time you go to bed each night. Your buddy lists and address books reveal your political affiliation and sexual orientation. Your e-mail headers reveal who is

which restaurants he eats at, whether he has a gym membership, and what nonprescription items he buys at a pharmacy. His phone reveals how often he goes to that gym, and his activity tracker reveals his activity level when he's there. Data from websites reveal what medical terms he's searched on. This is how a company like ExactData can sell lists of people who date online, people who gamble, and people who suffer from anxiety, incontinence, or erectile dysfunction.

PIERCING OUR ANONYMITY

When a powerful organization is eavesdropping on significant portions of our electronic infrastructure and can correlate the various surveillance streams, it can often identify people who are trying to hide. Here are four stories to illustrate that.

1. Chinese military hackers who were implicated in a broad set of attacks against the US government and corporations were identified because they accessed Facebook from the same network infrastructure they used to carry out their attacks.
2. Hector Monsegur, one of the leaders of the LulzSec hacker movement under investigation for breaking into numerous commercial networks, was identified and arrested in 2011 by the FBI. Although he usually practiced good computer security and used an anonymous relay service to protect his identity, he slipped up once. An inadvertent disclosure during a chat allowed an investigator to track down a video on YouTube of his car, then to find his Facebook page.
3. Paula Broadwell, who had an affair with CIA director David Petraeus, similarly took extensive precautions to hide her identity. She never logged in to her anonymous e-mail service from her home network. Instead, she used hotel and other public networks when she e-mailed him. The FBI correlated registration data from several different hotels—and hers was the common name.
4. A member of the hacker group Anonymous called "w0rmer," wanted for hacking US law enforcement websites, used an anonymous Twitter account, but linked to a photo of a woman's breasts taken with an iPhone. The photo's embedded GPS coordinates pointed to a house in Australia. Another website that referenced w0rmer also mentioned the name Higinio Ochoa. The police got hold of Ochoa's Facebook page, which included the information that he had an Australian girlfriend. Photos of the girlfriend matched the original photo that started all this, and police arrested w0rmer aka Ochoa.

Maintaining Internet anonymity against a ubiquitous surveillor is nearly impossible. If you forget even once to enable your protections, or click on the wrong link, or type the wrong thing, you've permanently attached your name to whatever anonymous provider you're using. The level of operational security required to maintain privacy and anonymity in the face of a focused and determined investigation is beyond the resources of even trained government agents. Even a team of highly trained Israeli assassins was quickly identified in Dubai, based on surveillance camera footage around the city.

The same is true for large sets of anonymous data. We might naïvely think that there are so many of us that it's easy to hide in the sea of data. Or that most of our data is anonymous. That's not true. Most techniques for anonymizing data don't work, and the data can be de-anonymized with surprisingly little information.

In 2006, AOL released three months of search data for 657,000 users: 20 million searches in all. The idea was that it would be useful for researchers; to protect people's identity, they replaced names with numbers. So, for example, Bruce Schneier might be 608429. They were surprised when researchers were able to attach names to numbers by correlating different items in individuals' search history.

In 2008, Netflix published 10 million movie rankings by 500,000 anonymized customers, as part of a challenge for people to come up with better recommendation systems than the one the company was using at that time. Researchers were able to de-anonymize people by comparing rankings and time stamps with public rankings and time stamps in the Internet Movie Database.

These might seem like special cases, but correlation opportunities pop up more frequently than you might think. Someone with access to an anonymous data set of telephone records, for example, might partially de-anonymize it by correlating it with a catalog merchant's telephone order database. Or Amazon's online book reviews could be the key to partially de-anonymizing a database of credit card purchase details.

Using public anonymous data from the 1990 census, computer scientist Latanya Sweeney found that 87% of the population in the United States, 216 million of 248 million people, could likely be uniquely identified by their five-digit ZIP code combined with their gender and date of birth. For about half, just a city, town, or municipality name was sufficient. Other researchers reported similar results using 2000 census data.

Google, with its database of users' Internet searches, could de-anonymize a public database of Internet purchases, or zero in on searches of medical terms to de-anonymize a public health database. Merchants who maintain detailed customer and purchase information could use their data to partially de-anonymize any large search engine's search data. A data broker holding databases of several companies might be able to de-anonymize most of the records in those databases.

Researchers have been able to identify people from their anonymous DNA by comparing the data with information from genealogy sites and other sources. Even something like Alfred Kinsey's sex research data from the 1930s and 1940s isn't safe. Kinsey took great pains to preserve the anonymity of his subjects, but in 2013, researcher Raquel Hill was able to identify 97% of them.

It's counterintuitive, but it takes less data to uniquely identify us than we think. Even though we're all pretty typical, we're nonetheless distinctive. It turns out that if you eliminate the top 100 movies everyone watches, our movie-watching habits are all pretty individual. This is also true for our book-reading habits, our Internet-shopping habits, our telephone habits, and our web-searching habits. We can be uniquely identified by our relationships. It's quite obvious that you can be uniquely identified by your location data. With 24/7 location data from your cell phone, your name can be uncovered without too much trouble. You don't even need all that data; 95% of Americans can be identified *by name* from just four time/date/location points.

The obvious countermeasures for this are, sadly, inadequate. Companies have anonymized data sets by removing some of the data, changing the time stamps, or inserting deliberate errors into the unique ID numbers they replaced names with. It turns out, though, that these sorts of tweaks only make de-anonymization slightly harder.

This is why regulation based on the concept of "personally identifying information" doesn't work. PII is usually defined as a name, unique account number, and so on, and

based on that information.

It's less Big Brother, and more hundreds of tattletale little brothers.

Today, Internet surveillance is far more insistent than cookies. In fact, there's a minor arms race going on. Your browser—yes, even Google Chrome—has extensive controls to block or delete cookies, and many people enable those features. DoNotTrackMe is one of the most popular browser plug-ins. The Internet surveillance industry has responded with "flash cookies"—basically, cookie-like files that are stored with Adobe's Flash player and remain when browsers delete their cookies. To block those, you can install FlashBlock. But there are other ways to uniquely track you, with esoteric names like evercookies, canvas fingerprinting, and cookie synching. It's not just marketers; in 2014, researchers found that the White House website used evercookies, in violation of its own privacy policy. I'll give some advice about blocking web surveillance in Chapter 15.

Cookies are inherently anonymous, but companies are increasingly able to correlate them with other information that positively identifies us. You identify yourself willingly to lots of Internet services. Often you do this with only a username, but increasingly usernames can be tied to your real name. Google tried to compel this with its "real name policy," which mandated that users register for Google Plus with their legal names, until it rescinded that policy in 2014. Facebook pretty much demands real names. Anytime you use your credit card number to buy something, your real identity is tied to any cookies set by companies involved in that transaction. And any browsing you do on your smartphone is tied to you as the phone's owner, although the website might not know it.

FREE AND CONVENIENT

Surveillance is the business model of the Internet for two primary reasons: people like free, and people like convenient. The truth is, though, that people aren't given much of a choice. It's either surveillance or nothing, and the surveillance is conveniently invisible so you don't have to think about it. And it's all possible because US law has failed to keep up with changes in business practices.

Before 1993, the Internet was entirely noncommercial, and free became the online norm. When commercial services first hit the Internet, there was a lot of talk about how to charge for them. It quickly became clear that, except for a few isolated circumstances like investment and porn websites, people weren't willing to pay even a small amount for access. Much like the business model for television, advertising was the only revenue model that made sense, and surveillance has made that advertising more profitable. Websites can charge higher prices for personally targeted advertising than they can for broadcast advertising. This is how we ended up with nominally free systems that collect and sell our data in exchange for services, then blast us with advertising.

"Free" is a special price, and there has been all sorts of psychological research showing that people don't act rationally around it. We overestimate the value of free. We consume more of something than we should when it's free. We pressure others to consume it. Free warps our normal sense of cost vs. benefit, and people end up trading their

Customer surveillance is much older than the Internet. Before the Internet, there were four basic surveillance streams. The first flowed from companies keeping records on their customers. This was a manufacturing supply company knowing what its corporate customers order, and who does the ordering. This was Nordstrom remembering its customers' sizes and the sorts of tailoring they like, and airlines and hotels keeping track of their frequent customers. Eventually this evolved into the databases that enable companies to track their sales leads all the way from initial inquiry to final purchase, and retail loyalty cards, which offer consumers discounts but whose real purpose is to track their purchases. Now lots of companies offer Customer Relationship Management, or CRM, systems to corporations of all sizes.

The second traditional surveillance stream was direct marketing. Paper mail was the medium, and the goal was to provide companies with lists of people who wanted to receive the marketing mail and not waste postage on people who did not. This was necessarily coarse, based on things like demographics, magazine subscriptions, or customer lists from related enterprises.

The third stream came from credit bureaus. These companies collected detailed credit information about people, and sold that information to banks trying to determine whether to give individuals loans and at what rates. This has always been a relatively expensive form of personal data collection, and only makes sense when lots of money is at stake: issuing credit cards, allowing someone to lease an apartment, and so on.

The fourth stream was from government. It consisted of various public records: birth and death certificates, driver's license records, voter registration records, various permits and licenses, and so on. Companies have increasingly been able to download, or purchase, this public data.

Credit bureaus and direct marketing companies combined these four streams to become modern day data brokers like Acxiom. These companies buy your personal data from companies you do business with, combine it with other information about you, and sell it to companies that want to know more about you. And they've ridden the tides of computerization. The more data you produce, the more they collect and the more accurately they profile you.

The breadth and depth of information that data brokers have is astonishing. They collect demographic information: names, addresses, telephone numbers, e-mail addresses, gender, age, marital status, presence and ages of children in household, education level, profession, income level, political affiliation, cars driven, and information about homes and other property. They collect lists of things you've purchased, when you've purchased them, and how you paid for them. They keep track of deaths, divorces, and diseases in your family. They collect everything about what you do on the Internet.

Data brokers use your data to sort you into various marketable categories. Want lists of people who fall into the category of "potential inheritor" or "adult with senior parent," or addresses of households with a "diabetic focus" or "senior needs"? Acxiom can provide you with that. InfoUSA has sold lists of "suffering seniors" and gullible seniors. In 2011, the data broker Teletrack sold lists of people who had applied for nontraditional credit

products like payday loans to companies who wanted to target them for bad financial deals. In 2012, the broker Equifax sold lists of people who were late on their mortgage payments to a discount loan company. Because this was financial information, both brokers were fined by the FTC for their actions. Almost everything else is fair game.

PERSONALIZED ADVERTISING

We use systems that spy on us in exchange for services. It's just the way the Internet works these days. If something is free, you're not the customer; you're the product. Or, as Al Gore said, "We have a stalker economy."

Advertising has always suffered from the problem that most people who see an advertisement don't care about the product. A beer ad is wasted on someone who doesn't drink beer. A car advertisement is largely wasted unless you are in the market for a car. But because it was impossible to target ads individually, companies did the best they could with the data they had. They segmented people geographically, and guessed which magazines and TV shows would best attract their potential customers. They tracked populations as a whole, or in large demographic groups. It was very inefficient. There's a famous quote, most reliably traced to the retail magnate John Wanamaker: "I know half of my advertising is wasted. The trouble is, I don't know which half."

Ubiquitous surveillance has the potential to change that. If you know exactly who wants to buy a lawn mower or is worried about erectile dysfunction, you can target your advertising to the right person at the right time, eliminating waste. (In fact, a national lawn care company uses aerial photography to better market its services.) And if you know the details about that potential customer—what sorts of arguments would be most persuasive, what sorts of images he finds most appealing—your advertising can be even more effective.

This also works in political advertising, and is already changing the way political campaigns are waged. Obama used big data and individual marketing to great effect in both 2008 and 2012, and other candidates across parties are following suit. This data is used to target fund-raising efforts and individualized political messages, and ensure that you actually get to the polls on Election Day—assuming the database says that you're voting for the correct candidate.

A lot of commercial surveillance data is filled with errors, but this information can be valuable even if it isn't very accurate. Even if you ended up showing your ad to the wrong people a third of the time, you could still have an effective advertising campaign. What's important is not perfect targeting accuracy; it's that the data is enormously better than before.

For example, in 2013, researchers were able to determine the physical locations of people on Twitter by analyzing similarities with other Twitter users. Their accuracy rate wasn't perfect—they were only able to predict a user's city with 58% accuracy—but for plenty of commercial advertising that level of precision is good enough.

Still, a lot of evidence suggests that surveillance-based advertising is oversold. There

music industry, Amazon versus the traditional publishing industry, Uber versus taxi companies. The new information middlemen are winning.

Google CEO Eric Schmidt said it: "We believe that modern technology platforms, such as Google, Facebook, Amazon and Apple, are even more powerful than most people realize … , and what gives them power is their ability to grow—specifically, their speed to scale. Almost nothing, short of a biological virus, can scale as quickly, efficiently or aggressively as these technology platforms and this makes the people who build, control, and use them powerful too."

What Schmidt is referring to is the inherently monopolistic nature of information middlemen. A variety of economic effects reward first movers, penalize latecomer competitors, entice people to join the largest networks, and make it hard for them to switch to a competing system. The result is that these new middlemen have more power than those they replaced.

Google controls two-thirds of the US search market. Almost three-quarters of all Internet users have Facebook accounts. Amazon controls about 30% of the US book market, and 70% of the e-book market. Comcast owns about 25% of the US broadband market. These companies have enormous power and control over us simply because of their economic position.

They all collect and use our data to increase their market dominance and profitability. When eBay first started, it was easy for buyers and sellers to communicate outside of the eBay system because people's e-mail addresses were largely public. In 2001, eBay started hiding e-mail addresses; in 2011, it banned e-mail addresses and links in listings; and in 2012, it banned them from user-to-user communications. All of these moves served to position eBay as a powerful intermediary by making it harder for buyers and sellers to take a relationship established inside of eBay and move it outside of eBay.

Increasingly, companies use their power to influence and manipulate their users. Websites that profit from advertising spend a lot of effort making sure you spend as much time on those sites as possible, optimizing their content for maximum addictiveness. The few sites that allow you to opt out of personalized advertising make that option difficult to find. Once companies combine these techniques with personal data, the result is going to be even more insidious.

Our relationship with many of the Internet companies we rely on is not a traditional company–customer relationship. That's primarily because we're not customers. We're products those companies sell to their *real* customers. The relationship is more feudal than commercial. The companies are analogous to feudal lords, and we are their vassals, peasants, and—on a bad day—serfs. We are tenant farmers for these companies, working on their land by producing data that they in turn sell for profit.

Yes, it's a metaphor—but it often really feels like that. Some people have pledged allegiance to Google. They have Gmail accounts, use Google Calendar and Google Docs, and have Android phones. Others have pledged similar allegiance to Apple. They have iMacs, iPhones, and iPads, and let iCloud automatically synchronize and back up

7

Political Liberty and Justice

In 2013, the First Unitarian Church of Los Angeles sued the NSA over its domestic spying, claiming that its surveillance of church members' telephone calling habits discouraged them from banding together to advocate for political causes. The church wasn't just being paranoid. In the 1950s and 1960s, the FBI monitored its minister because of his politics. Today, the church is worried that people, both Americans and foreigners, will end up on watch lists because of their association with the church.

Government surveillance is costly. Most obviously, it's extraordinarily expensive: $72 billion a year in the US. But it's also costly to our society, both domestically and internationally. Harvard law professor Yochai Benkler likens NSA surveillance to an autoimmune disease, because it attacks all of our other systems. It's a good analogy.

The biggest cost is liberty, and the risk is real enough that people across political ideologies are objecting to the sheer invasiveness and pervasiveness of the surveillance system. Even the politically conservative and probusiness *Economist* magazine argued, in a 2013 editorial about video surveillance, that it had gone too far: "This is where one of this newspaper's strongly held beliefs that technological progress should generally be welcomed, not feared, runs up against an even deeper impulse, in favour of liberty. Freedom has to include some right to privacy: if every move you make is being chronicled, liberty is curtailed."

ACCUSATION BY DATA

In the 17th century, the French statesman Cardinal Richelieu famously said, "Show me six lines written by the most honest man in the world, and I will find enough therein to hang him." Lavrentiy Beria, head of Joseph Stalin's secret police in the old Soviet Union, declared, "Show me the man, and I'll show you the crime." Both were saying the same thing: if you have enough data about someone, you can find sufficient evidence to find him guilty of *something*. It's the reason many countries' courts prohibit the police from engaging in "fishing expeditions." It's the reason the US Constitution specifically prohibits general warrants—documents that basically allow the police to search for *anything*. General warrants can be extremely abusive; they were used by the British in colonial America as a form of social control.

Ubiquitous surveillance means that anyone could be convicted of lawbreaking, once the police set their minds to it. It is incredibly dangerous to live in a world where everything you do can be stored and brought forward as evidence against you at some later date. There is significant danger in allowing the police to dig into these large data sets and find "evidence" of wrongdoing, especially in a country like the US with so many vague and punitive laws, which give prosecutors discretion over whom to charge with what, and

with overly broad material witness laws. This is especially true given the expansion of the legally loaded terms "terrorism," to include conventional criminals, and "weapons of mass destruction," to include almost anything, including a sawed-off shotgun. The US terminology is so broad that someone who donates $10 to Hamas's humanitarian arm could be considered a terrorist.

Surveillance puts us at risk of abuses by those in power, even if we're doing nothing wrong at the time of surveillance. The definition of "wrong" is often arbitrary, and can quickly change. For example, in the US in the 1930s, being a Communist or Socialist was a bit of an intellectual fad, and not considered wrong among the educated classes. In the 1950s, that changed dramatically with the witch-hunts of Senator Joseph McCarthy, when many intelligent, principled American citizens found their careers destroyed once their political history was publicly disclosed. Is someone's reading of Occupy, Tea Party, animal rights, or gun rights websites going to become evidence of subversion in five to ten years?

This situation is exacerbated by the fact that we are generating so much data and storing it indefinitely. Those fishing expeditions can go into the past, finding things you might have done 10, 15, or 20 years ago … and counting. Today's adults were able to move beyond their youthful indiscretions; today's young people will not have that freedom. Their entire histories will be on the permanent record.

Another harm of government surveillance is the way it leads to people's being categorized and discriminated against. George Washington University law professor Daniel Solove calls the situation Kafkaesque. So much of this data is collected and used in secret, and we have no right to refute or even see the evidence against us. This will intensify as systems start using surveillance data to make decisions automatically.

Surveillance data has been used to justify numerous penalties, from subjecting people to more intensive airport security to deporting them. In 2012, before his Los Angeles vacation, 26-year-old Irishman Leigh Van Bryan tweeted, "Free this week, for quick gossip/prep before I go and destroy America." The US government had been surveilling the entire Twitter feed. Agents picked up Bryan's message, correlated it with airplane passenger lists, and were waiting for him at the border when he arrived from Ireland. His comment wasn't serious, but he was questioned for five hours and then sent back home. We know that bomb jokes in airports can get you detained; now it seems that you have to be careful making even vague promises of international rowdiness anywhere on the Internet.

In 2013, a Hawaiian man posted a video on Facebook showing himself drinking and driving. Police arrested him for the crime; his defense was that it was a parody and that no actual alcohol was consumed on the video.

It's worse in the UK. There, people have been jailed because of a racist tweet or a tasteless Facebook post. And it's even more extreme in other countries, of course, where people are routinely arrested and tortured for things they've written online.

Most alarming of all, the US military targets drone strikes partly based on their targets'

data. There are two types of drone targeting. The first is "targeted killing," where a known individual is located by means of electronic or other surveillance. The second is "signature strikes," where unidentified individuals are targeted on the basis of their behavior and personal characteristics: their apparent ages and genders, their location, what they appear to be doing. At the peak of drone operations in Pakistan in 2009 and 2010, half of all kills were signature strikes. We don't have any information about how accurate the profiling was.

This is wrong. We should be free to talk with our friends, or send a text message to a family member, or read a book or article, without having to worry about how it would look to the government: our government today, our government in five or ten years, or some other government. We shouldn't have to worry about how our actions might be interpreted or misinterpreted, or how they could be used against us. We should not be subject to surveillance that is essentially indefinite.

GOVERNMENT CENSORSHIP

Freedom also depends on the free circulation of ideas. Government censorship, often enabled by surveillance, stifles them both.

China protects its citizens from the "dangers" of outside news and opinions on the Internet by something called the Golden Shield or, more commonly, the Great Firewall of China. It's a massive project that took eight years and cost $700 million to build, and its job is to censor the Internet. The goal is less to banish harmful ideas or squelch speech, and more to prevent effective organization. The firewall works pretty well; those with technical savvy can evade it, but it blocks the majority of China's population from finding all sorts of things, from information about the Dalai Lama to many Western search sites.

There's more government censorship on the Internet today than ever before. And it's not just politics. Countries censor websites because of their sexual nature, the religious views they espouse, their hosting of gambling platforms, and their promotion of drug use or illegal activity. The citizens of most Middle Eastern countries live under pervasive censorship. France, Germany, and Austria censor neo-Nazi content, including online auctions of Nazi memorabilia; other countries censor sites that incite violence. Vietnam's "Decree 72" prohibits people from discussing current affairs online. Many countries censor content that infringes on copyright. The UK censors pornography by default, although you can still opt out of the censorship. In 2010, the US censored WikiLeaks.

Most censorship is enforced by surveillance, which leads to self-censorship. If people know the government is watching everything they say, they are less likely to read or speak about forbidden topics. This is the point behind a 2014 Russian law requiring bloggers to register with the government. This is why the Great Firewall of China works so well as a censorship tool: it's not merely the technical capabilities of the firewall, but the threat that people trying to evade it will be discovered and reported by their fellow citizens. Those who do the reporting don't even necessarily agree with the government; they might face penalties of their own if they do not report. Internet companies in China often censor their users beyond what is officially required.

And the more severe the consequences of getting caught, the more excessively people self-censor.

CHILLING EFFECTS

Surveillance has a potentially enormous chilling effect on society. US Supreme Court Justice Sonia Sotomayor recognized this in her concurring opinion in a 2012 case about the FBI's installing a GPS tracker in someone's car. Her comments were much broader: "Awareness that the Government may be watching chills associational and expressive freedoms. And the Government's unrestrained power to assemble data that reveal private aspects of identity is susceptible to abuse. The net result is that GPS monitoring—by making available at a relatively low cost such a substantial quantity of intimate information about any person whom the Government, in its unfettered discretion, chooses to track—may 'alter the relationship between citizen and government in a way that is inimical to democratic society.' "

Columbia University law professor Eben Moglen wrote that "omnipresent invasive listening creates fear. And that fear is the enemy of reasoned, ordered liberty." Surveillance is a tactic of intimidation.

In the US, we already see the beginnings of this chilling effect. According to a Human Rights Watch report, journalists covering stories on the intelligence community, national security, and law enforcement have been significantly hampered by government surveillance. Sources are less likely to contact them, and they themselves are worried about being prosecuted. Human Rights Watch concludes that stories in the national interest that need to be reported don't get reported, and that the public is less informed as a result. That's the chilling effect right there.

Lawyers working on cases where there is some intelligence interest—foreign government clients, drugs, terrorism—are also affected. Like journalists, they worry that their conversations are monitored and that discussions with their clients will find their way into the prosecution's hands.

Post-9/11 surveillance has caused writers to self-censor. They avoid writing about and researching certain subjects; they're careful about communicating with sources, colleagues, or friends abroad. A Pew Research Center study conducted just after the first Snowden articles were published found that people didn't want to talk about the NSA online. A broader Harris poll found that nearly half of Americans have changed what they research, talk about, and write about because of NSA surveillance. Surveillance has chilled Internet use by Muslim Americans, and by groups like environmentalists, gun-rights activists, drug policy advocates, and human rights workers. After the Snowden revelations of 2013, people across the world were less likely to search personally sensitive terms on Google.

A 2014 report from the UN High Commissioner on Human Rights noted, "Even the mere possibility of communications information being captured creates an interference with privacy, with a potential chilling effect on rights, including those to free expression

we learned that the NSA had been spying on UN communications, in violation of international law. We know that there have been all sorts of other abuses of state and local government surveillance authorities as well.

Abuses happen inside surveillance organizations as well. For example, NSA employees routinely listen to personal phone calls of Americans overseas, and intercept e-mail and pass racy photos around the office. We learned this from two intercept operators in 2008, and again from Snowden in 2014. We learned from the NSA that its agents sometimes spy on people they know; internally, they call this practice LOVEINT. The NSA's own audit documents note that the agency broke its own privacy rules 2,776 times in 12 months, from 2011 to 2012. That's a lot—eight times a day—but the real number is probably much higher. Because of how the NSA polices itself, it essentially decides how many violations it discovers.

This is not a new problem, nor one limited to the NSA. Recent US history illustrates many episodes in which surveillance has been systematically abused: against labor organizers and suspected Communists after World War I, against civil rights leaders, and against Vietnam War protesters. The specifics aren't pretty, but it's worth giving a couple of them.

- Through extensive surveillance, J. Edgar Hoover learned of Martin Luther King's extramarital affairs, and in an anonymous letter he tried to induce him to commit suicide in 1964: "King, look into your heart. You know you are a complete fraud and a great liability to all of us Negroes. White people in this country have enough frauds of their own but I am sure they don't have one at this time anywhere near your equal. You are no clergyman and you know it. I repeat you are a colossal fraud and an evil, vicious one at that. You could not believe in God… . Clearly you don't believe in any personal moral principles… . King, there is only one thing left for you to do. You know what it is. You have just 34 days in which to do it (this exact number has been selected for a specific reason, it has definite practical significance). You are done. There is but one way out for you. You better take it before your filthy, abnormal fraudulent self is bared to the nation."
- This is how a Senate investigation described the FBI's COINTELPRO surveillance program in 1976: "While the declared purposes of these programs were to protect the 'national security' or prevent violence, Bureau witnesses admit that many of the targets were nonviolent and most had no connections with a foreign power. Indeed, nonviolent organizations and individuals were targeted because the Bureau believed they represented a 'potential' for violence—and nonviolent citizens who were against the war in Vietnam were targeted because they gave 'aid and comfort' to violent demonstrators by lending respectability to their cause… . But COINTELPRO was more than simply violating the law or the Constitution. In COINTELPRO the Bureau secretly took the law into its own hands, going beyond the collection of intelligence and beyond its law enforcement function to act outside the legal process altogether and to covertly disrupt, discredit and harass groups and individuals."

Nothing has changed. Since 9/11, the US has spied on the Occupy movement, pro- and anti-abortion activists, peace activists, and other political protesters.

- The NSA and FBI spied on many prominent Muslim Americans who had nothing to do with terrorism, including Faisal Gill, a longtime Republican Party operative and onetime candidate for public office who held a top-secret security clearance and served in the Department of Homeland Security under President George W. Bush; Asim Ghafoor, a prominent attorney who has represented clients in terrorism-related cases; Hooshang Amirahmadi, an Iranian American professor of international relations at Rutgers University; and Nihad Awad, the executive director of the largest Muslim civil rights organization in the country.
- The New York Police Department went undercover into minority neighborhoods. It monitored mosques, infiltrated student and political groups, and spied on entire communities. Again, people were targeted because of their ethnicity, not because of any accusations of crimes or evidence of wrongdoing. Many of these operations were conducted with the help of the CIA, which is prohibited by law from spying on Americans.

There's plenty more. Boston's fusion center spied on Veterans for Peace, the women's

8

Commercial Fairness and Equality

Accretive Health is a debt collection agency that worked for a number of hospitals in Minnesota. It was in charge of billing and collection for those hospitals, but it also coordinated scheduling, admissions, care plans, and duration of hospital stays. If this sounds like a potential conflict of interest, it was. The agency collected extensive patient data and used it for its own purposes, without disclosing to patients the nature of its involvement in their healthcare. It used information about patient debts when scheduling treatment and harassed patients for money in emergency rooms. The company denied all wrongdoing, but in 2012 settled a Minnesota lawsuit by agreeing not to operate in Minnesota for two to six years. On the one hand, the fact that Accretive was caught and punished shows that the system is working. On the other hand, it also shows how easy it is for our data to be mishandled and misused.

Stories like this demonstrate the considerable risk to society in allowing corporations to conduct mass surveillance. It's their surveillance that contributes to all of the offenses against civil liberties, social progress, and freedom that I described in the previous chapter. And in addition to enabling government surveillance, corporate surveillance carries its own risks.

SURVEILLANCE-BASED DISCRIMINATION

In a fundamental way, companies use surveillance data to discriminate. They place people into different categories and market goods and services to them differently on the basis of those categories.

"Redlining" is a term from the 1960s to describe a practice that's much older: banks discriminating against members of minority groups when they tried to purchase homes. Banks would not approve mortgages in minority neighborhoods—they would draw a red line on their maps delineating those zones. Or they would issue mortgages to minorities only if they were buying houses in predominantly minority neighborhoods. It's illegal, of course, but for a long time banks got away with it. More generally, redlining is the practice of denying or charging more for services by using neighborhood as a proxy for race—and it's much easier to do on the Internet.

In 2000, Wells Fargo bank created a website to promote its home mortgages. The site featured a "community calculator" to help potential buyers search for neighborhoods. The calculator collected the current ZIP code of the potential customers and steered them to neighborhoods based on the predominant race of that ZIP code. The site referred white residents to white neighborhoods, and black residents to black neighborhoods.

This practice is called weblining, and it has the potential to be much more pervasive

and much more discriminatory than traditional redlining. Because corporations collect so much data about us and can compile such detailed profiles, they can influence us in many different ways. A 2014 White House report on big data concluded, "… big data analytics have the potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, employment, health, education, and the marketplace." I think the report understated the risk.

Price discrimination is also a big deal these days. It's not discrimination in the same classic racial or gender sense as weblining; it's companies charging different people different prices to realize as much profit as possible. We're most familiar with this concept with respect to airline tickets. Prices change all the time, and depend on factors like how far in advance we purchase, what days we're traveling, and how full the flight is. The airline's goal is to sell tickets to vacationers at the bargain prices they're willing to pay, while at the same time extracting from business travelers the much higher amounts that *they're* willing to pay. There is nothing nefarious about the practice; it's just a way of maximizing revenues and profits. Even so, price discrimination can be *very* unpopular. Raising the price of snow shovels after a snowstorm, for example, is considered price-gouging. This is why it is often cloaked in things like special offers, coupons, or rebates.

Some types of price discrimination are illegal. For example, a restaurant cannot charge different prices depending on the gender or race of the customer. But it can charge different prices based on time of day, which is why you see lunch and dinner menus with the same items and different prices. Offering senior discounts and special children's menus is legal price discrimination. Uber's surge pricing is also legal.

In many industries, the options you're offered, the price you pay, and the service you receive depend on information about you: bank loans, auto insurance, credit cards, and so on. Internet surveillance facilitates a fine-tuning of this practice. Online merchants already show you different prices and options based on your history and what they know about you. Depending on who you are, you might see a picture of a red convertible or a picture of a minivan in online car ads, and be offered different options for financing and discounting when you visit dealer websites. According to a 2010 *Wall Street Journal* article, the price you pay on the Staples website depends on where you are located, and how close a competitor's store is to you. The article states that other companies, like Rosetta Stone and Home Depot, are also adjusting prices on the basis of information about the individual user.

More broadly, we all have a customer score. Data brokers assign it to us. It's like a credit score, but it's not a single number, and it's focused on what you buy, based on things like purchasing data from retailers, personal financial information, survey data, warranty card registrations, social media interactions, loyalty card data, public records, website interactions, charity donor lists, online and offline subscriptions, and health and fitness information. All of this is used to determine what ads and offers you see when you browse the Internet.

In 2011, the US Army created a series of recruiting ads showing soldiers of different genders and racial backgrounds. It partnered with a cable company to deliver those ads

increase in how often you see that car. This is, essentially, the business model of search engines. In their early days, there was talk about how an advertiser could pay for better placement in search results. After public outcry and subsequent guidance from the FTC, search engines visually differentiated between "natural" results by algorithm and paid ones. So now you get paid search results in Google framed in yellow, and paid search results in Bing framed in pale blue. This worked for a while, but recently the trend has shifted back. Google is now accepting money to insert particular URLs into search results, and not just in the separate advertising areas. We don't know how extensive this is, but the FTC is again taking an interest.

When you're scrolling through your Facebook feed, you don't see every post by every friend; what you see has been selected by an automatic algorithm that's not made public. But people can pay to increase the likelihood that their friends or fans will see their posts. Payments for placement represent a significant portion of Facebook's income. Similarly, a lot of those links to additional articles at the bottom of news pages are paid placements.

The potential for manipulation here is enormous. Here's one example. During the 2012 election, Facebook users had the opportunity to post an "I Voted" icon, much like the real stickers many of us get at polling places after voting. There is a documented bandwagon effect with respect to voting; you are more likely to vote if you believe your friends are voting, too. This manipulation had the effect of increasing voter turnout 0.4% nationwide. So far, so good. But now imagine if Facebook manipulated the visibility of the "I Voted" icon on the basis of either party affiliation or some decent proxy of it: ZIP code of residence, blogs linked to, URLs liked, and so on. It didn't, but if it had, it would have had the effect of increasing voter turnout in one direction. It would be hard to detect, and it wouldn't even be illegal. Facebook could easily tilt a close election by selectively manipulating what posts its users see. Google might do something similar with its search results.

A truly sinister social networking platform could manipulate public opinion even more effectively. By amplifying the voices of people it agrees with, and dampening those of people it disagrees with, it could profoundly distort public discourse. China does this with its 50 Cent Party: people hired by the government to post comments on social networking sites supporting, and to challenge comments opposing, party positions. Samsung has done much the same thing.

Many companies manipulate what you see according to your user profile: Google search, Yahoo News, even online newspapers like the *New York Times*. This is a big deal. The first listing in a Google search result gets a third of the clicks, and if you're not on the first page, you might as well not exist. The result is that the Internet you see is increasingly tailored to what your profile indicates your interests are. This leads to a phenomenon that political activist Eli Pariser has called the "filter bubble": an Internet optimized to your preferences, where you never have to encounter an opinion you don't agree with. You might think that's not too bad, but on a large scale it's harmful. We don't want to live in a society where everybody only ever reads things that reinforce their existing opinions, where we never have spontaneous encounters that enliven, confound,

9

Business Competitiveness

In 1993, the Internet was a very different place from what it is today. There was no electronic commerce; the World Wide Web was in its infancy. The Internet was a communications tool for techies and academics, and we used e-mail, newsgroups, and a chat protocol called IRC. Computers were primitive, as was computer security. For about 20 years, the NSA had managed to keep cryptography software out of the mainstream by classifying it as a munition and restricting its export. US products with strong cryptography couldn't be sold overseas, which meant that US hardware and software companies put weak—and by that I mean easily breakable—cryptography into both their domestic and their international products, because that was easier than maintaining two versions.

But the world was changing. Cryptographic discoveries couldn't be quashed, and the academic world was catching up to the capabilities of the NSA. In 1993, I wrote my first book, *Applied Cryptography*, which made those discoveries accessible to a more general audience. It was a big deal, and I sold 180,000 copies in two editions. *Wired* magazine called it "the book the National Security Agency wanted never to be published," because it taught cryptographic expertise to non-experts. Research was international, and non-US companies started springing up, offering strong cryptography in their products. One study from 1993 found over 250 cryptography products made and marketed outside the US. US companies feared that they wouldn't be able to compete, because of the export restrictions in force.

At the same time, the FBI started to worry that strong cryptography would make it harder for the bureau to eavesdrop on the conversations of criminals. It was concerned about e-mail, but it was most concerned about voice encryption boxes that could be attached to telephones. This was the first time the FBI used the term "going dark" to describe its imagined future of ubiquitous encryption. It was a scare story with no justification to support it, just as it is today—but lawmakers believed it. They passed the CALEA law I mentioned in Chapter 6, and the FBI pushed for them to ban all cryptography without a backdoor.

Instead, the Clinton administration came up with a solution: the Clipper Chip. It was a system of encryption with surveillance capabilities for FBI and NSA access built in. The encryption algorithm was alleged to be strong enough to prevent eavesdropping, but there was a backdoor that allowed someone who knew the special key to get at the plaintext. This was marketed as "key escrow" and was billed as a great compromise; trusted US companies could compete in the world market with strong encryption, and the FBI and NSA could maintain their eavesdropping capabilities.

The first encryption device with the Clipper Chip built in was an AT&T secure phone.

It wasn't a cell phone; this was 1993. It was a small box that plugged in between the wired telephone and the wired handset and encrypted the voice conversation. For the time, it was kind of neat. The voice quality was only okay, but it worked.

No one bought it.

In retrospect, it was rather obvious. Nobody wanted encryption with a US government backdoor built in. Privacy-minded individuals didn't want it. US companies didn't want it. And people outside the US didn't want it, especially when there were non-US alternatives available with strong cryptography and no backdoors. The US government was the only major customer for the AT&T devices, and most of those were never even used.

Over the next few years, the government tried other key escrow initiatives, all designed to give the US government backdoor access to all encryption, but the market soundly rejected all of those as well.

The demise of the Clipper Chip, and key escrow in general, heralded the death of US government restrictions on strong cryptography. Export controls were gradually lifted over the next few years, first on software in 1996 and then on most hardware a few years later. The change came not a moment too soon. By 1999, over 800 encryption products from 35 countries other than the US had filled the market.

What killed both the Clipper Chip and crypto export controls were not demands for privacy from consumers. Rather, they were killed by the threat of foreign competition and demands from US industry. Electronic commerce needed strong cryptography, and even the FBI and the NSA could not stop its development and adoption.

GOVERNMENT SURVEILLANCE COSTS BUSINESS

Those of us who fought the crypto wars, as we call them, thought we had won them in the 1990s. What the Snowden documents have shown us is that instead of dropping the notion of getting backdoor government access, the NSA and FBI just kept doing it in secret. Now that this has become public, US companies are losing business overseas because their non-US customers don't want their data collected by the US government.

NSA surveillance is costing US companies business in three different ways: people fleeing US cloud providers, people not buying US computer and networking equipment, and people not trusting US companies.

When the story about the NSA's getting user data directly from US cloud providers—the PRISM program—broke in 2013, businesses involved faced a severe public relations backlash. Almost immediately, articles appeared noting that US cloud companies were losing business and their counterparts in countries perceived as neutral, such as Switzerland, were gaining. One survey of British and Canadian companies from 2014 found that 25% of them were moving their data outside the US, even if it meant decreased performance. Another survey of companies found that NSA revelations made executives much more concerned about where their data was being stored.

Estimates of how much business will be lost by US cloud providers vary. One 2013

10

Privacy

The most common misconception about privacy is that it's about having something to hide. "If you aren't doing anything wrong, then you have nothing to hide," the saying goes, with the obvious implication that privacy only aids wrongdoers.

If you think about it, though, this makes no sense. We do nothing wrong when we make love, go to the bathroom, or sing in the shower. We do nothing wrong when we search for a job without telling our current employer. We do nothing wrong when we seek out private places for reflection or conversation, when we choose not to talk about something emotional or personal, when we use envelopes for our mail, or when we confide in a friend and no one else.

Moreover, even those who say that don't really believe it. In a 2009 interview, Google CEO Eric Schmidt put it this way: "If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place." But in 2005, Schmidt banned employees from talking to reporters at CNET because a reporter disclosed personal details about Schmidt in an article. Facebook's Mark Zuckerberg declared in 2010 that privacy is no longer a "social norm," but bought the four houses abutting his Palo Alto home to help ensure his own privacy.

There are few secrets we don't tell *someone*, and we continue to believe something is private even after we've told that person. We write intimate letters to lovers and friends, talk to our doctors about things we wouldn't tell anyone else, and say things in business meetings we wouldn't say in public. We use pseudonyms to separate our professional selves from our personal selves, or to safely try out something new.

Facebook's CEO Mark Zuckerberg showed a remarkable naïveté when he stated, "You have one identity. The days of you having a different image for your work friends or co-workers and for the other people you know are probably coming to an end pretty quickly. Having two identities for yourself is an example of a lack of integrity."

We're not the same to everyone we know and meet. We act differently when we're with our families, our friends, our work colleagues, and so on. We have different table manners at home and at a restaurant. We tell different stories to our children than to our drinking buddies. It's not necessarily that we're lying, although sometimes we do; it's that we reveal different facets of ourselves to different people. This is something innately human. Privacy is what allows us to act appropriately in whatever setting we find ourselves. In the privacy of our home or bedroom, we can relax in a way that we can't when someone else is around.

Privacy is an inherent human right, and a requirement for maintaining the human condition with dignity and respect. It is about choice, and having the power to control how

you present yourself to the world. Internet ethnographer danah boyd puts it this way: "Privacy doesn't just depend on agency; being able to achieve privacy is an expression *of* agency."

When we lose privacy, we lose control of how we present ourselves. We lose control when something we say on Facebook to one group of people gets accidentally shared with another, and we lose complete control when our data is collected by the government. "How did he know that?" we ask. How did I lose control of who knows about my traumatic childhood, my penchant for tasteless humor, or my vacation to the Dominican Republic? You may know this feeling: you felt it when your mother friended you on Facebook, or on any other social networking site that used to be just you and your friends. Privacy violations are intrusions.

There's a strong physiological basis for privacy. Biologist Peter Watts makes the point that a desire for privacy is innate: mammals in particular don't respond well to surveillance. We consider it a physical threat, because animals in the natural world are surveilled by predators. Surveillance makes us feel like prey, just as it makes the surveillors act like predators.

Psychologists, sociologists, philosophers, novelists, and technologists have all written about the effects of constant surveillance, or even just the perception of constant surveillance. Studies show that we are less healthy, both physically and emotionally. We have feelings of low self-esteem, depression, and anxiety. Surveillance strips us of our dignity. It threatens our very selves as individuals. It's a dehumanizing tactic employed in prisons and detention camps around the world.

Violations of privacy are not all equal. Context matters. There's a difference between a Transportation Security Administration (TSA) officer finding porn in your suitcase and your spouse finding it. There's a difference between the police learning about your drug use and your friends learning about it. And violations of privacy aren't all equally damaging. Those of us in marginal socioeconomic situations—and marginalized racial, political, ethnic, and religious groups—are affected more. Those of us in powerful positions who are subject to people's continued approval are affected more. The lives of some of us depend on privacy.

Our privacy is under assault from constant surveillance. Understanding how this occurs is critical to understanding what's at stake.

THE EPHEMERAL

Through most of history, our interactions and conversations have been ephemeral. It's the way we naturally think about conversation. Exceptions were rare enough to be noteworthy: a preserved diary, a stenographer transcribing a courtroom proceeding, a political candidate making a recorded speech.

This has changed. Companies have fewer face-to-face meetings. Friends socialize online. My wife and I have intimate conversations by text message. We all behave as if these conversations were ephemeral, but they're not. They're saved in ways we have no

14

Solutions for Corporations

As we look to limit corporate surveillance, it's important to remember that we all reap enormous benefits from data collection and use. Data collection gives us many benefits and conveniences that just weren't possible before: real-time driving directions based on actual congestion data, grocery lists that remember what we bought last time, the ability to get a store refund even if you don't save your receipts, the ability to remotely verify that you turned out the lights and locked the door, instant communication with anyone anywhere in the world. There's more coming. Watch any science fiction movie or television show and pay attention to the marvels of a fully computerized world; much of it assumes that computers know, respond to, and remember what people are doing. This sort of surveillance is our future, and it's a future filled with things that make our lives better and more enjoyable.

Similarly, there is value to unfettered access to technology. Although much of this book focuses on the dark side of technology, we must remember that technology has been an enormous benefit to us all. Technology enables us to accomplish complex tasks more quickly, easily, and accurately for many purposes: to develop more durable construction materials; to find and disseminate information; to precisely depict physical phenomena; to communicate with others free of geographical constraints; to document events; to grow more food; to live longer. I could not have written this book without the Internet. It's not perfect, of course. Technology is unevenly distributed on the planet, and there are haves and have-nots, but—in general—more technology is better.

The last thing we want to do is derail that future. We simply don't know what sorts of inventions are coming, or what major human problems they will be able to solve. We need to be able to experiment with new technologies and with new businesses based on those technologies, and this includes surveillance technologies. The trick will be maximizing the benefits that come from companies collecting, storing, and analyzing our data, while minimizing the harms.

There are lots of solutions out there to consider. The 1980 OECD Privacy Framework is a great place to start; it lays out limitations on data collection, data storage, and data use. In 1995, the European Union passed the EU Data Protection Directive, which regulated personal data collected by corporations. American corporations, accustomed to the much more permissive legal regime in the US, are constantly running afoul of European law. And reforms, bringing that law up to date with modern technology, are currently being debated.

The solutions offered in this chapter are all directed at the private collection and use of our data. Sometimes these changes can be spurred by the market, but most of the time they will be facilitated by laws. This is really a list of things governments need to do, which in

turn is really a list of things citizens need to demand that their governments do. Since they affect corporations, they're in this chapter.

MAKE VENDORS LIABLE FOR PRIVACY BREACHES

One way to improve the security of collected data is to make companies liable for data breaches.

Corporations are continually balancing costs and benefits. In this case, the costs consist of the cost of securing the data they collect and save, the cost of insecurity if there's a breach, and the value of the data they collect. Right now, the cost of insecurity is low. A few very public breaches aside—Target is an example here—corporations find it cheaper to spend money on PR campaigns touting good security, weather the occasional press storm and round of lawsuits when they're proven wrong, and fix problems after they become public.

OECD Privacy Framework (1980)

COLLECTION LIMITATION PRINCIPLE: There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

DATA QUALITY PRINCIPLE: Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

PURPOSE SPECIFICATION PRINCIPLE: The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

USE LIMITATION PRINCIPLE: Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except: a) with the consent of the data subject; or b) by the authority of law.

SECURITY SAFEGUARDS PRINCIPLE: Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.

OPENNESS PRINCIPLE: There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

INDIVIDUAL PARTICIPATION PRINCIPLE: Individuals should have the right: a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to them; b) to have communicated to them, data relating to them i. within a reasonable time; ii. at a charge, if any, that is not excessive; iii. in a reasonable manner; and iv. in a form that is readily intelligible to them; c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and d) to challenge data relating to them and, if the challenge is successful to have the data erased, rectified, completed or amended.

ACCOUNTABILITY PRINCIPLE: A data controller should be accountable for complying with measures which give effect to the principles stated above.

This is because most of the cost of privacy breaches falls on the people whose data is exposed. In economics, this is known as an externality: an effect of a decision not borne by the decision maker. Externalities limit the incentive for companies to improve their

collected about us is collected because we want it to be. We object when that information is being used in ways we didn't intend: when it is stored, shared, sold, correlated, and used to manipulate us in some stealthy way. This means that we need restrictions on how our data can be used, especially restrictions on ways that differ from the purposes for which it was collected.

Other problems arise when corporations treat their underlying algorithms as trade secrets: Google's PageRank algorithm, which determines what search results you see, and credit-scoring systems are two examples. The companies have legitimate concerns about secrecy. They're worried both that competitors will copy them and that people will figure out how to game them. But I believe transparency trumps proprietary claims when the algorithms have a direct impact on the public. Many more algorithms can be made public—or redesigned so they can be made public—than currently are. For years, truth in lending and fair lending laws have required financial institutions to ensure that the algorithms they use are both explainable and legally defensible. Mandated transparency needs to be extended into other areas where algorithms hold power over people: they have to be open. Also, there are ways of auditing algorithms for fairness without making them public.

Corporations tend to be rational risk assessors, and will abide by regulation. The key to making this work is oversight and accountability. This isn't something unusual: there are many regulated industries in our society, because we know what they do is both important and dangerous. Personal information and the algorithms used to analyze it are no different. Some regular audit mechanism would ensure that corporations are following the rules, and would penalize them if they don't.

This all makes sense in theory, but actually doing it is hard. The last thing we want is for the government to start saying, "You can only do this and nothing more" with our data. Permissions-based regulation would stifle technological innovation and change. We want rights-based regulation—basically, "You can do anything you want unless it is prohibited."

REGULATE DATA COLLECTION AS WELL

Regulating data use isn't enough. Privacy needs to be regulated in many places: at collection, during storage, upon use, during disputes. The OECD Privacy Framework sets them out nicely, and they're all essential.

There's been a concerted multi-year effort by US corporations to convince the world that we don't need regulations on data collection, only on data use. Companies seek to eradicate any limitations on data collection because they know that any use limitations will be narrowly defined, and that they can slowly expand them once they have our data. (A common argument against any particular data-use regulation is that it's a form of censorship.) They know that if collection limitations are in place, it's much harder to change them. But as with government mass surveillance, the privacy harms come from the simple collection of the data, not only from its use. Remember the discussion of algorithmic surveillance from Chapter 10. Unrestricted corporate collection will result in broad collection, expansive sharing with the government, and a slow chipping away at the

necessarily narrowly defined use restrictions.

We need to fight this campaign. Limitations on data collection aren't new. Prospective employers are not allowed to ask job applicants whether they're pregnant. Loan applications are not allowed to ask about the applicants' race. "Ban the Box" is a campaign to make it illegal for employers to ask about applicants' criminal pasts. The former US gays-in-the-military compromise, "Don't Ask Don't Tell," was a restriction on data collection. There are restrictions on what questions can be asked by the US Census Bureau.

Extending this to a world where everything we do is mediated by computers isn't going to be easy, but we need to start discussing what sorts of data should never be collected. There are some obvious places to start. What we read online should be as private as it is in the paper world. This means we should legally limit recording the webpages we read, the links we click on, and our search results. It's the same with our movements; it should not be a condition of having a cell phone that we subject ourselves to constant surveillance. Our associations—to whom we communicate, whom we meet on the street—should not be continually monitored. Maybe companies can be allowed to use some of this data immediately and then must purge it. Maybe they'll be allowed to save it for a short period of time.

One intriguing idea has been proposed by University of Miami Law School professor Michael Froomkin: requiring both government agencies and private companies engaging in mass data collection to file Privacy Impact Notices, modeled after Environmental Impact Reports. This would serve to inform the public about what's being collected and why, and how it's being stored and used. It would encourage decision makers to think about privacy early in any project's development, and to solicit public feedback.

One place to start is to require opt-in. Basically, there are two ways to obtain consent. Opt-in means that you have to explicitly consent before your data is collected and used. Opt-out is the opposite; your data will be collected and used unless you explicitly object. Companies like Facebook prefer opt-out, because they can make the option difficult to find and know that most people won't bother. Opt-in is more fair, and the use of service shouldn't be contingent on allowing data collection.

Right now, there's no downside to collecting and saving everything. By limiting what companies can collect and what they can do with the data they collect, by making companies responsible for the data under their control, and by forcing them to come clean with customers about what they actually collect and what they do with it, we will influence them to collect and save only the data about us they know is valuable.

Congress needs to begin the hard work of updating US privacy laws and stop making excuses for inaction. Courts can also play a significant role safeguarding consumer privacy by enforcing current privacy laws. The regulatory agencies, such as the FTC and the FCC, have some authority to protect consumer privacy in certain domains. But what the United States needs today is an independent data protection agency comparable to those in other countries around the world. And we have to do better than patching problems only after they become sufficiently harmful. These challenges are big and

able to go to any data broker and say, "I'm not your product. I never gave you permission to gather information about me and sell it to others. I want my data out of your database." This is what the EU is currently grappling with: the right to be forgotten. In 2014, the European Court of Justice ruled that in some cases search engines need to remove information about individuals from their results. This caused a torrent of people demanding that Google remove search results that reflected poorly on them: politicians, doctors, pedophiles. We can argue about the particulars of the case, and whether the court got the balance right, but this is an important right for citizens to have with respect to their data that corporations are profiting from.

US Consumer Privacy Bill of Rights (2012)

INDIVIDUAL CONTROL: Consumers have a right to exercise control over what personal data companies collect from them and how they use it.

TRANSPARENCY: Consumers have a right to easily understandable and accessible information about privacy and security practices.

RESPECT FOR CONTEXT: Consumers have a right to expect that companies will collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data.

SECURITY: Consumers have a right to secure and responsible handling of personal data.

ACCESS AND ACCURACY: Consumers have a right to access and correct personal data in usable formats, in a manner that is appropriate to the sensitivity of the data and the risk of adverse consequences to consumers if the data is inaccurate.

FOCUSED COLLECTION: Consumers have a right to reasonable limits on the personal data that companies collect and retain.

ACCOUNTABILITY: Consumers have a right to have personal data handled by companies with appropriate measures in place to assure they adhere to the Consumer Privacy Bill of Rights.

MAKE DATA COLLECTION AND PRIVACY SALIENT

We reveal data about ourselves all the time, to family, friends, acquaintances, lovers, even strangers. We share with our doctors, our investment counselors, our psychologists. We share a lot of data. But we think of that sharing transactionally: I'm sharing data with *you*, because I need you to know things/trust you with my secrets/am reciprocating because you've just told me something personal.

As a species, we have evolved all sorts of psychological systems to navigate these complex privacy decisions. And these systems are extraordinarily complex, highly attuned, and delicately social. You can walk into a party and immediately know how to behave. Whom you talk to, what you tell to whom, who's around you, who's listening: most of us can navigate that beautifully. The problem is that technology inhibits that social ability. Move that same party onto Facebook, and suddenly our intuition starts failing. We forget who's reading our posts. We accidentally send something private to a public forum. We don't understand how our data is monitored in the background. We don't realize what the technologies we're using can and cannot do.

In large part that's because the degree of privacy in online environments isn't salient. Intuition fails when thoughts of privacy fade into the background. Once we can't directly

perceive people, we don't do so well. We don't think, "There's a for-profit corporation recording everything I say and trying to turn that into advertising." We don't think, "The US and maybe other governments are recording everything I say and trying to find terrorists, or criminals, or drug dealers, or whoever is the bad guy this month." That's not what's obvious. What's obvious is, "I'm at this virtual party, with my friends and acquaintances, and we're talking about personal stuff."

And so we can't use people's continual exposure of their private data on these sites as evidence of their consent to be monitored. What they're consenting to is the real-world analogue they have in their heads, and they don't fully understand the ramifications of moving that system into cyberspace.

Companies like Facebook prefer it this way. They go out of their way to make sure you're not thinking about privacy when you're on their site, and they use cognitive tricks like showing you pictures of your friends to increase your trust. Governments go even further, making much of their surveillance secret so people don't even know it's happening. This explains the disconnect between people's claims that privacy is important and their actions demonstrating that it isn't: the systems we use are deliberately designed so that privacy issues don't arise.

We need to give people the option of true online privacy, and the ability to understand and choose that option. Companies will be less inclined to do creepy things with our data if they have to justify themselves to their customers and users. And users will be less likely to be seduced by "free" if they know the true costs. This is going to require "truth in product" laws that will regulate corporations, and similar laws to regulate government.

For starters, websites should be required to disclose what third parties are tracking their visitors, and smartphone apps should disclose what information they are recording about their users. There are too many places where surveillance is hidden; we need to make it salient as well.

Again, this is hard. Notice, choice, and consent is the proper way to manage this, but we know that lengthy privacy policies written in legalese—those notice-and-consent user agreements you click "I agree" to without ever reading—don't work. They're deliberately long and detailed, and therefore boring and confusing; and they don't result in any meaningful consent on the part of the user. We can be pretty sure that a pop-up window every time you post something to Facebook saying, "What you've written will be saved by Facebook and used for marketing, and will be given to the government on demand," won't work, either. We need some middle way. My guess is that it will involve standardized policies and some sort of third-party certification.

ESTABLISH INFORMATION FIDUCIARIES

In several areas of our lives we routinely give professionals access to very personal information about ourselves. To ensure that they only use that information in our interests, we have established the notion of fiduciary responsibility. Doctors, lawyers, and accountants are all bound by rules that require them to put the interests of their clients

we're going to have to learn to live with that.

The problem is that we're too good at adapting, at least in the short term. People who grow up with more surveillance will be more comfortable with it. Some of us went to schools with ID checks and metal detectors. Some of us work at buildings that demand daily badge checks. Most of us who fly in the US now accept TSA searches. And all of us who shop are no longer surprised about massive thefts of credit card numbers. These are all ways in which we have accustomed ourselves to having less privacy. Like many fundamental rights, privacy is one of those things that becomes noticed only when it's gone. That's unfortunate, because after it's gone it's much harder to regain.

We have to stop the slide. Fundamentally, the argument for privacy is a moral one. It is something we ought to have—not because it is profitable or efficient, but because it is moral. Mass surveillance should be relegated to the dustbin of history, along with so many other practices that humans once considered normal but are now universally recognized as abhorrent. Privacy is a human right. This isn't a new idea. Privacy is recognized as a fundamental right in the Universal Declaration of Human Rights (1948) and the European Convention on Human Rights (1970).

It's in the US Constitution—not explicitly, but it's implied in the Fourth, Fifth, and Ninth Amendments. It's part of the 2000 Charter of Fundamental Rights of the European Union. In 2013, the UN General Assembly approved a resolution titled "The right to privacy in the digital age," affirming that our fundamental right to privacy applies online as well as offline, and the risk of mass surveillance undermines this right.

The principles are enshrined in both national and international law. We need to start following them. Privacy is not a luxury that we can only afford in times of safety. Instead, it's a value to be preserved. It's essential for liberty, autonomy, and human dignity. We must understand that privacy is not something to be traded away in some fearful attempt to guarantee security, but something to maintain and protect in order to have real security.

Charter of Fundamental Rights of the European Union (2000)

ARTICLE 7: Respect for private and family life. Everyone has the right to respect for his or her private and family life, home and communications.

ARTICLE 8: Protection of personal data. 1. Everyone has the right to the protection of personal data concerning him or her. 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. 3. Compliance with these rules shall be subject to control by an independent authority.

None of this will happen without a change of attitude. In the end, we'll get the privacy we as a society demand and not a bit more.

DON'T WAIT

The longer we wait to make changes, the harder it will become. On the corporate side, ubiquitous tracking and personalized advertising are already the norm, and companies have strong lobbying presences to stymie any attempt to change that. California's Do Not

As individuals and as a society, we are constantly trying to balance our different values. We never get it completely right. What's important is that we deliberately engage in the process. Too often the balancing is done for us by governments and corporations with their own agendas.

Whatever our politics, we need to get involved. We don't want the FBI and NSA to secretly decide what levels of government surveillance are the default on our cell phones; we want Congress to decide matters like these in an open and public debate. We don't want the governments of China and Russia to decide what censorship capabilities are built into the Internet; we want an international standards body to make those decisions. We don't want Facebook to decide the extent of privacy we enjoy amongst our friends; we want to decide for ourselves. All of these decisions are bigger and more important than any one organization. They need to be made by a greater and more representative and inclusive institution. We want the public to be able to have open debates about these things, and "we the people" to be able to hold decision makers accountable.

I often turn to a statement by Rev. Martin Luther King Jr: "The arc of history is long, but it bends toward justice." I am long-term optimistic, even if I remain short-term pessimistic. I think we will overcome our fears, learn how to value our privacy, and put rules in place to reap the benefits of big data while securing ourselves from some of the risks. Right now, we're seeing the beginnings of a very powerful worldwide movement to recognize privacy as a fundamental human right, not just in the abstract sense we see in so many public pronouncements, but in a meaningful and enforceable way. The EU is leading the charge, but others will follow. The process will take years, possibly decades, but I believe that in half a century people will look at the data practices of today the same way we now view archaic business practices like tenant farming, child labor, and company stores. They'll look immoral. The start of this movement, more than anything else, will be Edward Snowden's legacy.

I started this book by talking about data as exhaust: something we all produce as we go about our information-age business. I think I can take that analogy one step further. Data is the pollution problem of the information age, and protecting privacy is the environmental challenge. Almost all computers produce personal information. It stays around, festering. How we deal with it—how we contain it and how we dispose of it—is central to the health of our information economy. Just as we look back today at the early decades of the industrial age and wonder how our ancestors could have ignored pollution in their rush to build an industrial world, our grandchildren will look back at us during these early decades of the information age and judge us on how we addressed the challenge of data collection and misuse.

We should try to make them proud.